

## Interest-Driven Model for Human Dynamics \*

SHANG Ming-Sheng(尚明生)<sup>1</sup>, CHEN Guan-Xiong(陈冠雄)<sup>2</sup>, DAI Shuang-Xing(戴双星)<sup>2</sup>,  
WANG Bing-Hong(汪秉宏)<sup>2,3</sup>, ZHOU Tao(周涛)<sup>1,2,4\*\*</sup>

<sup>1</sup>Web Sciences Center, School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054

<sup>2</sup>Department of Modern Physics, University of Science and Technology of China, Hefei 230026

<sup>3</sup>The Research Center for Complex System Science, University of Shanghai for Science and Technology, Shanghai 200093

<sup>4</sup>Department of Physics, University of Fribourg, Chemin du Musée 3, 1700 Fribourg, Switzerland

*Empirical observations indicate that the interevent time distribution of human actions exhibits heavy-tailed features. The queuing model based on task priorities is to some extent successful in explaining the origin of such heavy tails, however, it cannot explain all the temporal statistics of human behavior especially for the daily entertainments. We propose an interest-driven model, which can reproduce the power-law distribution of interevent time. The exponent can be analytically obtained and is in good accordance with the simulations. This model well explains the observed relationship between activities and power-law exponents, as reported recently for web-based behavior and the instant message communications.*

Due to the complexity and diversity of our behavior, in the literature, the human individual activity pattern is usually simplified as a completely random point-process, which can be well described by the Poisson process, leading to an exponential interevent time distribution. That is to say, the time difference between two consecutive events should be almost uniform, and long gaps are hardly observed. However, recent empirical studies have shown us a completely different picture: in both group level<sup>[1–8]</sup> and individual level,<sup>[9–18]</sup> the interevent time distributions display a heavy-tailed nature, and usually can be approximately fitted by power laws. These empirical findings lead to a new research direction, namely, human dynamics,<sup>[19,20]</sup> which have attracted increasing attention recently.

Based on the queuing theory, Barabasi *et al.*<sup>[10,12,21]</sup> proposed a simple model where the individual executes the highest-priority task first, and they suggested that the highest-priority-first (HPF) protocol is a potential origin of those heavy tails. The queuing model has great success in explaining the heavy tails in many human-oriented dynamics. However, some other human activity patterns, also displaying the similar heavy-tailed phenomenon, could not be explained by a task-based mechanism. For example, the web browsing,<sup>[13,16]</sup> activities in online societies,<sup>[14,18]</sup> and playing on-line games<sup>[4]</sup> are mainly driven by personal interests, which could not be treated as tasks needing to be executed. The in-depth understanding of the non-Poisson statistics in those interest-driven

systems requires a new model out of the perspective of the queuing theory. Han *et al.*<sup>[22]</sup> proposed a model based on adaptive human interest, which can reproduce the power-law interevent time distribution with exponent  $-1$ . However, real human-initiated systems are of various exponents.<sup>[20]</sup> In this Letter, we propose a new interest-based model for human dynamics, which can reproduce the power-law distribution of interevent time with the tunable exponent. The exponent can be analytically obtained and is in good accordance with the simulations. This model well explains the observed relationship between activities and power-law exponents, as reported recently for online activities<sup>[13,14,16,18]</sup> and the instant message communications.<sup>[17]</sup>

Considering the daily experience: when we come across a new thing, we may be zesty to do it again but after a period not doing so, the interest may be weakened. Then, for a long rest, an occasional action may remind us the good feeling, and the new cycle starts. Accordingly, we propose a simple model as follows: (i) the interest at time step  $t$  is quantified by the probability that an action will occur in this time step and denoted by  $x(t)$ ; (ii) at each time step  $t$ , if an action occurs (with probability  $x(t)$ ), then we reset the interest as  $x(t) = 1$ ; (iii) at each time step  $t$ , if  $t'$  is the last time step an action occurs, the interest is set as

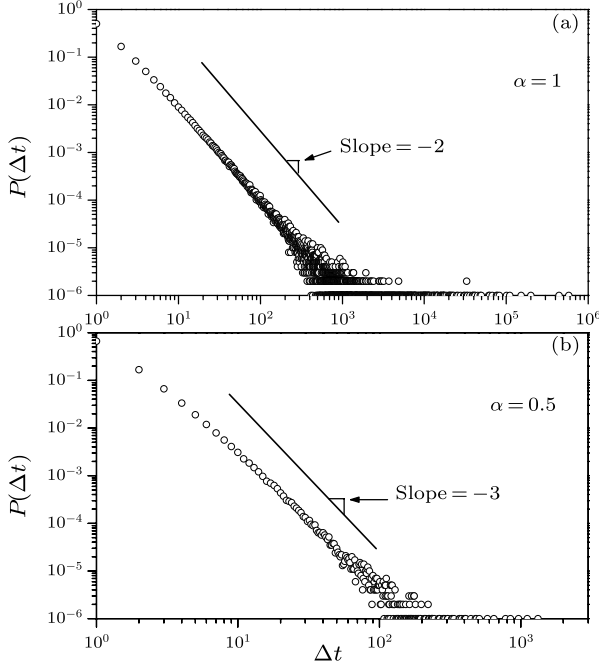
$$x(t) = \frac{1}{1 + \alpha(t - t')}, \quad (1)$$

where  $\alpha > 0$  is a free parameter. Clearly, if an action

\*Supported by the National Natural Science Foundation of China under Grant Nos 70871082, 10975126, 90924011, 70971089, 10635040 and 60973069, the China Postdoctoral Science Foundation under Grant No 20080431273, and the Sino-Swiss Science and Technology Cooperation (SSSTC) Project (EG 20-032009).

\*\*Email: zhutou@ustc.edu

occurs at time step  $t$ , the probability another action happens again at time step  $t + 1$  is  $x(t + 1) = \frac{1}{1+\alpha}$ . Initially, we set  $x(0) = 1$ .



**Fig. 1.** Simulation results of interevent time distribution. To mimic the real situation (in the individual, no averaging data is available), we only perform one-shot simulation with  $10^6$  time intervals. The slopes of the two solid lines are the analytical result  $1 + \frac{1}{\alpha}$ , while the fitting exponents, obtained by using the maximum likelihood estimation,<sup>[23,24]</sup> are 2.02 for  $\alpha = 1$  and 2.91 for  $\alpha = 0.5$ , respectively. The numerical results are in good accordance with the analytical ones.

If at a certain time step  $t$ , an action occurs, the probability the next action occurs at time step  $t + \Delta t$  ( $\Delta t = 1, 2, 3, \dots$ ) is

$$P(\Delta t) = \frac{1}{1 + \alpha \Delta t} \prod_{i=1}^{\Delta t-1} \left(1 - \frac{1}{1 + \alpha i}\right). \quad (2)$$

It straightforwardly arrives at

$$P(\Delta t) = \frac{1}{\alpha} \frac{(\Delta t - 1)! \alpha^{\Delta t}}{\prod_{i=1}^{\Delta t} (1 + \alpha i)}. \quad (3)$$

Dividing by  $\alpha^{\Delta t}$  simultaneously for the numerator and denominator, one obtains

$$P(\Delta t) = \frac{1}{\alpha} \frac{(\Delta t - 1)!}{\prod_{i=1}^{\Delta t} (i + \frac{1}{\alpha})}. \quad (4)$$

Using the approximation

$$\prod_{i=1}^{\Delta t} \left(i + \frac{1}{\alpha}\right) \approx \frac{\Gamma(\Delta t + 1 + \frac{1}{\alpha})}{\Gamma(1 + \frac{1}{\alpha})}, \quad (5)$$

where

$$\Gamma(z) = \int_0^\infty y^{z-1} e^{-y} dy, \quad (6)$$

and if  $z$  is a positive integer,  $\Gamma(z) = (z - 1)!$ . According to Eqs. (5) and (6), Eq. (4) can be rewritten as

$$\begin{aligned} P(\Delta t) &= \frac{1}{\alpha} \frac{\Gamma(\Delta t) \Gamma(1 + \frac{1}{\alpha})}{\Gamma(\Delta t + 1 + \frac{1}{\alpha})} \\ &= \frac{1}{\alpha} B\left(\Delta t, 1 + \frac{1}{\alpha}\right) \sim \frac{1}{\alpha} \Delta t^{-(1 + \frac{1}{\alpha})}, \end{aligned} \quad (7)$$

where

$$B(m, n) = \frac{\Gamma(m) \Gamma(n)}{\Gamma(m + n)}. \quad (8)$$

Accordingly, the interevent time distribution approximately follows a power law with exponent depending on the parameter  $\alpha$ . The analytical results are in good accordance with the simulations. Two examples are reported in Fig. 1, where the parameter is respectively set as  $\alpha = 1$  and  $\alpha = 0.5$ . The exponents obtain by simulations and the analytical result agree well with each other.

In summary, we have proposed an interest-driven model that can produce the power-law interevent time distribution. A considerable advantage compared with the model introduced by Han *et al.*<sup>[22]</sup> is that the power-law exponent in the present model is tunable. Although the mechanism underlying this model is very simple, it has rich implication for many real observations. As reported very recently by Zhou *et al.*,<sup>[14,17]</sup> the activity of a user plays a major role in determining his interevent time distribution: in the statistical level, the higher activity corresponds to larger power-law exponent. This phenomenon challenges both the queuing model<sup>[10,12,21]</sup> that leads to two universal classes with exponent equal to 1 and 1.5 respectively and the adaptive model based on human interest<sup>[22]</sup> which gives a fixed exponent 1. Note that the average interevent time can be estimated as

$$\overline{\Delta t} = \int_1^\infty \Delta t P(\Delta t) d\Delta t. \quad (9)$$

According to Eq. (7), this integration convergent only if  $\alpha < 1$ , and then we have

$$\overline{\Delta t} \sim \frac{1}{1 - \alpha}. \quad (10)$$

Since the average activity is inverse to the average interevent time, it is easily observed that the present model can reproduce the phenomenon that statistically a more active individual will have larger exponent. The present model suggests the memory effect to be the origin of the heavy-tailed temporal statistics. Clearly, if an event occurs recently, an event will very possibly occur at this moment; while if during a long time no event occurs, it will less possibly occur at time moment. This is a strong short-range memory effect that has already been demonstrated by a number of real systems.<sup>[25,26]</sup> In addition, the present model is

indeed a cyclic dynamics and does not have any long-range memory effect since when an event occurs, everything goes back to the initial state. Therefore, the present indicates that in the individual level, the significant long-range temporal correlation may not exist, whose validity needs further empirical exploration. The cyclic dynamics as well as the specific decaying function of interest with time are ideal and rough approximation of the real human behavior, which may actually be far different from individual to individual. In contrast to the oversimplified mechanism underlying the present model, we hope that some insights of human dynamics can be obtained therein.

We acknowledge Xiaopu Han and Wei Hong for their useful discussions.

## References

- [1] Caldarelli G, Marsili M and Zhang Y C 1997 *Europhys. Lett.* **40** 479
- [2] Mainardi F, Raberto M, Gorenflo R and Scalas E 2000 *Physica A* **287** 468
- [3] Plerou V, Gopikrishnan P, Amaral L A N, Gabaix X and Stanley H E 2000 *Phys. Rev. E* **62** 3023
- [4] Henderson T and Nhatti S 2001 *Proceedings of the 9th ACM International Conference on Multimedia* (New York: ACM Press) p 212
- [5] Masoliver J, Montero M and Weiss G H 2003 *Phys. Rev. E* **67** 021112
- [6] Scalas E, Gorenflo R, Luckock H, Mainardi H, Mantelli M and Raberto M 2004 *Quant. Finance* **4** 695
- [7] Kaizoji T and Kaizoji M 2004 *Physica A* **336** 563
- [8] Scalas E, Kaizoji T, Kirchler M, Huber J and Tedeschi A 2006 *Physica A* **366** 463
- [9] Eckmann J P, Moses E and Sergi D 2004 *Proc. Natl. Acad. Sci. U.S.A.* **101** 14333
- [10] Barabási A L 2005 *Nature* **435** 207
- [11] Oliveira J G and Barabási A L 2005 *Nature* **437** 1251
- [12] Vázquez A, Oliveira J G, Dezső Z, Goh K I, Kondor I and Barabási A L 2006 *Phys. Rev. E* **73** 036127
- [13] Dezső Z, Almaas E, Lukács A, Rácz B, Szakadát I and Barabási A L 2006 *Phys. Rev. E* **73** 066132
- [14] Zhou T, Kiet H A T, Kim B J, Wang B H and Holme P 2008 *Europhys. Lett.* **82** 28002
- [15] Li N N, Zhang N and Zhou T 2008 *Physica A* **387** 6391
- [16] Gonçalves B and Ramasco J J 2008 *Phys. Rev. E* **78** 026123
- [17] Hong W, Han X P, Zhou T and Wang B H 2009 *Chin. Phys. Lett.* **26** 028902
- [18] Radicchi F 2009 *Phys. Rev. E* **80** 026118
- [19] Barabási A L 2007 *IEEE Control Systems Magazine* **27** 33
- [20] Zhou T, Han X P and Wang B H 2008 *Science Matters: Humanities as Complex Systems* ed Burguete M and Lam L (Singapore: World Scientific Publishing) p 207
- [21] Vázquez A 2005 *Phys. Rev. Lett.* **95** 248701
- [22] Han X P, Zhou T and Wang B H 2008 *New J. Phys.* **10** 073010
- [23] Goldstein M L, Morris S A and Yen G G 2004 *Eur. Phys. J. B* **41** 255
- [24] Newman M E J 2005 *Contemp. Phys.* **46** 323
- [25] Vázquez A 2007 *Physica A* **373** 747
- [26] Goh K I and Barabasi A L 2008 *Europhys. Lett.* **81** 48002